ARTICLES

Lazzarini's Lucky Approximation of π

LEE BADGER Weber State University Ogden, UT 84408-1702

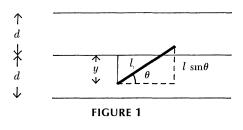
1. Introduction

In 1812 Laplace [14] remarked that one could approximate π by performing a Buffon needle experiment. Since then several needle casters claim to have done just that. Lazzarini's 1901 Buffon approximation of π [15] was accurate to six decimal places. This work was commended in several publications for illustrating the connectedness of mathematics [13] and validating the laws of probability [3], [7]. However, the 1960 study of Gridgeman [9] suggested that Lazzarini's experiment was not carried out in an entirely legitimate fashion and perhaps didn't warrant the praise it later received. But Gridgeman stopped short of establishing that the experiment was contrived to achieve the desired numerical result.

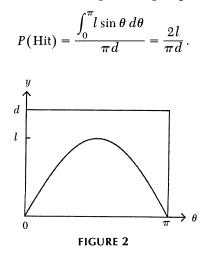
I will begin by reviewing the history of Lazzarini's experiment and the work of Gridgeman that debunked it. I will then extend Gridgeman's work to virtually rule out any possibility that Lazzarini performed a valid experiment. Some of this work was anticipated by that of O'Beirne [16], however, it also goes beyond that of O'Beirne. In this study elementary applications of probability, recurrence relations, and various numerical techniques are used to look deeper into a small piece of the history of mathematics. For brief expositions of the work of Gridgeman and O'Beirne see also Pilton [17] and Zaydel [18].

2. Buffon's Needle

In 1777 Georges-Louis Leclerc, Comte de Buffon, published the results of an earlier study that has come to be known as the Buffon Needle Problem [4]. In its simplest form it assumes that a needle of length l is cast at random on an infinite plane, ruled with parallel lines of uniform separation d where d > l. It asks for the probability of the event that the needle intersects one of the lines. Buffon found this probability to be $2l/\pi d$. At this point let us review the technique of solution.



Assume the grid is oriented in FIGURE 1, y measures the perpendicular distance from the lower end of the needle to the nearest grid line above it, and θ measures the smallest counterclockwise angle from the grid direction to the needle. There is clearly a one-to-one correspondence between possible tosses of the needle and ordered pairs (θ, y) , where $0 \le \theta < \pi$ and $0 \le y < d$. The needle hits one of the grid lines if, and only if, $y < l \sin \theta$. A random toss of the needle means that the needle's vertical displacement (y) and orientation (θ) with respect to the grid are each random and uniformly distributed and so the probability of a hit is the ratio of the area under the curve $y = l \sin \theta$ to the area of the rectangular sample space of FIGURE 2, that is



Elementary properties alone imply that the probability of a hit is proportional to the length of the needle. This observation opens an interesting "back door" method to obtain the Buffon result (see Gnedenko [8]). Suppose a convex polygon with n sides of length l_1, l_2, \ldots, l_n is tossed at random onto the grid. Also suppose the polygon is of diameter less than d and that a hit occurs if, and only if, exactly two sides hit. Then

$$P(\text{polygon hits}) = \frac{1}{2} \sum_{i=1}^{n} k l_i$$
$$= ks/2 \quad \text{where } s \text{ is the perimeter.}$$

A limiting argument yields the same result for any closed convex curve.

Now let's apply the result when the curve is a circle of radius r with 2r < d. Looking at where the center of the circle falls at random between grid lines we see that

$$P(\text{circle hits}) = 2r/d, \text{ so}$$
$$k(2\pi r)/2 = 2r/d$$

and hence $k = 2/(\pi d)$. So

$$P(i \text{th side hits}) = \sum_{j \neq i} p_{ij} = kl_i$$
$$= \frac{2l_i}{\pi d},$$

which is the original Buffon result.

It is intuitively clear that the probability should be an increasing function of l and a decreasing function of d, but that the probability depends on π is perhaps unexpected and has been a source of many other studies—including this one.

3. Lazzarini's "Approximation" of π

Laplace recognized that Buffon's result could be used to obtain an experimental approximation of π . If one casts N needles and if H of them hit, then since the theoretical probability of a hit is approximated by the relative frequency, $2l/\pi d \approx H/N$. It follows that

$$\hat{\pi} := \frac{2l}{d} \frac{N}{H} \approx \pi$$
 (A := B means A is defined by B).

Several needle casters actually performed this experiment and published their experimental values of π . These results are summarized in Gridgeman. The focus of this paper is the experiment reported by Mario Lazzarini in 1901. One of Lazzarini's results has been widely quoted; in it l = 2.5 cm, d = 3 cm, N = 3408, and H = 1808, so that $\hat{\pi} = (2(2.5)/3)(3408/1808) = 3.1415929...$ Since $\pi = 3.1415926...$, Lazzarini's result gives six-decimal place accuracy.

But something seems a little suspect about those numbers 3408 and 1808. Why cast 3408 needles? Why not a nice round number like 1000 or 3500? Our skepticism increases if we look at what happens when the number of hits is increased or decreased by 1. If H = 1807, $\hat{\pi} = 3.1433...$; if H = 1809, $\hat{\pi} = 3.1398...$ Lazzarini appears to have been extraordinarily lucky!

Repeated trials are capable of a simple statistical analysis. If, using Lazzarini's grid and needle, one wants to be 95% confident that $|\pi - \hat{\pi}| < 0.5 \times 10^{-6}$ (six-decimal place accuracy), then one needs to cast around 134 trillion needles! To obtain this number, we seek N such that

$$P\left(\left| \, \pi - \frac{5N}{3H} \right| < 0.5 \times 10^{-6} \right) \ge 0.95.$$

Now $\left| \pi - \frac{5N}{3H} \right| < \epsilon$ if, and only if, $\left| \frac{3\pi}{5} - \frac{N}{H} \right| < \frac{3\epsilon}{5}$ if, and only if, (assuming 1/(x) is locally linear near $3\pi/5$) $\left| \frac{5}{3\pi} - \frac{H}{N} \right| < \frac{1}{(3\pi/5)^2} \frac{3\epsilon}{5} = \frac{5\epsilon}{3\pi^2}$ if, and only if, $\left| \frac{5N}{3\pi} - H \right| < \frac{5N\epsilon}{3\pi^2}$. Since *H* is binomially distributed with parameters *N* and $p = \frac{5}{3\pi}$, its expectation is *Np* and its variance is Np(1-p). Using the normal approximation, in order to have $P\left(\left| \frac{5N}{3\pi} - H \right| < \frac{5N}{3\pi^2} (0.5) \times 10^{-6} \right) = .95$ we need

$$\frac{\frac{3\pi^2}{3\pi^2}(0.5) \times 10^{-6}}{\sqrt{Np(1-p)}} \approx 1.96$$

or $N \approx 134 \times 10^{12}$.

In addition to questioning the number of needles cast, one may also question the accuracy of the measurement of l and d. Lazzarini reported that they were measured to be 2.5 cm and 3 cm, but gave no tolerances on his measuring instruments. At the turn of the century, state of the art micrometers had errors of about ± 0.0005 cm [12]. Incorporating these best error bounds, one calculates that $3.1404 < \hat{\pi} < 3.1427$. So the last four figures of agreement of Lazzarini's $\hat{\pi}$ with π are meaningless. Zaydel [18] gives additional analyses of measurement error in needle experiments.

Another source of skepticism emerges when we look more closely at Lazzarini's

$$\hat{\pi} = \frac{2(2.5)}{3} \frac{3408}{1808} = \frac{355}{113}.$$

This fraction is known to number theorists as a convergent in the continued fraction for π ([1] and [11]) and historians recognize it as a rational approximation to π discovered by the fifth century A.D. Chinese mathematician Tsu Ch'ung-chih [5]. So to many mathematicians—and presumably to Lazzarini—it was a well-known rational approximation of π . Also, in terms of the magnitude of its denominator, it is an exceedingly accurate approximation of π . The next smallest denominator that yields a strictly better approximation occurs in the fraction 52,163/16,604.

4. A Lesson in "Experimental Design"

Let us consider how we would go about rigging up a good Buffon experimental approximation to π . To get Lazzarini's approximation, we need to choose l, d, N, and H such that

$$\frac{2lN}{dH} = \frac{355}{113} = \frac{355k}{113k} = \frac{5 \cdot 71k}{113k}.$$

A reasonable choice might be 2l = 5 so l = 2.5 and d > l, say d = 3, resulting in (N/H) = (213k/113k). The net effect is that if, at any multiple of 213 casts, we have the same multiple of 113 hits, we achieve the desired approximation. Lazzarini achieved this at the sixteenth multiple.

In his conclusion Gridgeman suggested that the mysterious 3408 was selected as a stopping point only because it was a potential generator of 355/113 and that by the "very happiest of coincidences" the optimum H was observed. He went on to question whether Lazzarini performed any experiment at all, or if the results were purely mental concoctions.

I will give a more definitive answer to this question. Is there any chance that Lazzarini actually performed an experiment? Assuming that $\hat{\pi}$ is computed after each cast, one could stop at any point at which $\hat{\pi} = 355/113$. What is the likelihood of a sequence of casts yielding $\hat{\pi} = 355/113$ at *some* stopping point, for his choice of d = 3 and l = 2.5?

Let A_k denote the event that 113k hits occur during 213k casts, and let $a_k := P(A_k)$. If N needles are cast then there are [N/213] =: m such events. We are interested in the value of $P(A_1 \cup A_2 \cup \cdots \cup A_m) =: P(U_m)$, the probability that the ratio 113/213 is achieved at least once in the first 213m casts. We will find out in Section 6 that Lazzarini claimed to have dropped N = 4000 needles, which corresponds to m = 18. We will show shortly that $P(U_{18})$ is about 0.30. So there is a good possibility that Lazzarini could actually have performed his experiment and achieved his reported result at some point.

To get this value for $P(U_{18})$ we first use Stirling's approximation [6] on the binomial probabilities a_k and simplify:

$$a_{k} = \frac{(213k)!}{(113k)!(100k)!} p^{113k} (1-p)^{100k} \sim c \alpha^{k} / \sqrt{k}$$

where

$$\begin{aligned} \alpha &\coloneqq \frac{213^{213}}{113^{113}100^{100}} p^{113} (1-p)^{100} = 0.99999\,99999\,99132 + \\ c &\coloneqq \sqrt{\frac{213}{113 \cdot 2 \cdot 100\pi}} = 0.05477 + \text{ and } p = \frac{5}{3\pi}, \end{aligned}$$

and where we use "~" in the standard sense that the ratio of the two sides approaches 1 as $k \to \infty$. The error bounds of Stirling's formula [18] can be used to show that a_k is actually within 1% of this estimate for all k. An explanation for α 's closeness to 1 is that, considered as a function of p, α takes its maximum value of 1 when p = 113/213 and $5/3\pi$ is extremely close to 113/213.

Whenever one of the events A_i occurs, we may think of the entire experiment and the accounting of hits as beginning anew. In other words, for j > i, $P(A_j|A_i) = P(A_{j-i})$ and so $P(A_j \cap A_i) = a_i a_{j-i}$. Properties of this type will play an important role in our computations. For instance

$$\sum_{1 \le i < j < k \le m} P(A_i \cap A_j \cap A_k) = \sum_{i=1}^{m-2} \sum_{1 \le j < k \le m-i} P(A_i \cap A_{i+j} \cap A_{i+k})$$
$$= \sum_{i=1}^{m-2} \sum_{1 \le j < k \le m-i} P(A_i) P(A_{i+j} \cap A_{i+k} | A_i)$$
$$= \sum_{i=1}^{m-2} \sum_{1 \le j < k \le m-i} P(A_i) P(A_j \cap A_k)$$
$$= \sum_{i=1}^{m-2} a_i S(2, m-i)$$

where S(u, v) is the sum of the probabilities of all intersections of u members of A_1, \ldots, A_v . With $S(0, v) \coloneqq 1$ for $0 \le v \le m$, we can show by induction on u that $S(u, v) = \sum_{i=1}^{v-u+1} a_i S(u-1, v-i)$ for $v = u, u+1, \ldots, m$. By the principle of inclusion-exclusion, $P(U_m) = S(1, m) - S(2, m) + \cdots + (-1)^{m+1}S(m, m)$. This provides a reasonably efficient, $O(m^3/3)$, way to compute $P(U_m)$. Using this algorithm, we obtain $P(U_{18}) \approx 0.3041$.

The sequence $u_m := P(U_m)$ is slow to converge; $u_1 \approx 0.05$, $u_{10} \approx 0.23$, $u_{30} \approx 0.45$, $u_{100} \approx 0.55$, and $u_{500} \approx 0.76$; it is unclear what its limit is.

5. Is Ultimate Success a Certainty?

The next question is whether, in an *infinite* sequence of casts, the ratio 355/113 is certain to occur. That is, is $\lim_{m\to\infty} P(U_m) = 1$? The answer is no, and we can argue as follows. The expected number of occurrences of this ratio, $\sum_{k=1}^{\infty} a_k$ is finite because $a_k \sim c \alpha^k / \sqrt{k}$ and $\alpha < 1$. If the desired ratio occurs with probability 1, then it must occur with probability 1 at some finite A_k and then accounting of successes and failures could begin anew and we may argue that with probability 1 it must occur at some subsequent finite A_k and then at yet another and so on, ad infinitum. But since the expected number of such occurrences is finite, this is impossible.

More precisely, let B_n be the event "*n* or more A's". Then $B_1 = \bigcup_{k=1}^{\infty} A_k$. The sequence B_n is nested decreasing so $B_{n+1} = B_{n+1} \cap B_n$ and $P(B_{n+1}) = P(B_{n+1} \cap B_n) = P(B_n) \cdot P(B_{n+1}|B_n) = P(B_n) \cdot P(B_1)$, and by induction we have $P(B_n) = (P(B_1))^n$. If $P(B_1)$ were equal to 1, then we would have $P(B_n) = 1$ for all *n* and hence $P(\bigcap_{n=1}^{\infty} B_n) = \lim_{n \to \infty} P(B_n) = 1$. But $\bigcap_{n=1}^{\infty} B_n =$ "infinitely many A's" = $\bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k$. So $1 \le P(\bigcup_{k=j}^{\infty} A_k) \le \sum_{k=j}^{\infty} P(A_k)$ for all *j*, contradicting the convergence of $\sum_{k=1}^{\infty} a_k$. Thus $P(B_1) < 1$.

A more detailed analysis will yield a numerical approximation of $f := P(\bigcup_{k=1}^{\infty} A_k) = P(B_1)$, the probability of ultimate success. The value of f is related to

that of $a := \sum_{k=1}^{\infty} a_k$. The relation f = a/(1+a) is standard renewal theory and is derived by Feller [6] using generating functions. It can also be derived from the above analysis and the identity, $E(X) = \sum_{n=1}^{\infty} P(X \ge n)$ where X := "the number of A_k that occur." The following is a more direct, elementary proof.

Let F_n be the event "355/113 occurs first at trial 213*n*", i.e. $F_n = A_n \cap A_{n-1}^c \cap \cdots \cap A_1^c$. The *F*'s are mutually exclusive and $A_n = F_1 \cup \cdots \cup F_n$. So $P(\bigcup_{n=1}^{\infty} A_n) = f = P(\bigcup_{n=1}^{\infty} F_n) = \sum_{n=1}^{\infty} f_n$ where $f_n := P(F_n)$.

Also for $k \ge 1$, $a_k = P(A_k)$

$$= P(A_k \cap F_1) + P(A_k \cap F_2) + \dots + P(A_k \cap F_k)$$
(*)
$$= P(F_1)P(A_k|F_1) + P(F_2)P(A_k|F_2) + \dots + P(F_k)P(A_k|F_k)$$

$$= f_1a_{k-1} + f_2a_{k-2} + \dots + f_ka_0, \text{ where } a_0 = 1.$$

Expanding, collecting terms with like sums of indices, and using (*) one obtains

$$(f_1 + f_2 + \dots + f_n)(a_0 + a_1 + a_2 + \dots + a_n)$$

= $a_1 + a_2 + \dots + a_n + b_{n+1} + \dots + b_{2n}$, where $0 < b_k < a_k$.

So

$$\sum_{k=1}^{n} a_k < \sum_{k=1}^{n} f_k \cdot \sum_{k=0}^{n} a_k < \sum_{k=1}^{2n} a_k \text{ and}$$
$$\sum_{k=1}^{n} a_k / \left(1 + \sum_{k=1}^{n} a_k \right) < \sum_{k=1}^{n} f_k < \sum_{k=1}^{2n} a_k / \left(1 + \sum_{k=1}^{n} a_k \right) \text{ for } n = 1, 2, \dots$$

Letting $n \to \infty$ we obtain that $a < \infty$ implies f < 1, as obtained earlier, but furthermore, that f = a/(1+a) for all extended real a.

So to estimate f, we first need to estimate a and its Stirling approximation, $\sum_{n=1}^{\infty} c \alpha^n / \sqrt{n}$.

Since α^x / \sqrt{x} is decreasing for x > 0,

$$\sum_{n=1}^{\infty} \alpha^n / \sqrt{n} > \int_1^{\infty} \alpha^x / \sqrt{x} \, dx$$

and

$$\sum_{n=1}^{\infty} \alpha^n / \sqrt{n} = \alpha + \sum_{n=2}^{\infty} \alpha^n / \sqrt{n} < \alpha + \int_1^{\infty} \alpha^x / \sqrt{x} \, dx.$$

By the change of variable $t := \sqrt{-\ln \alpha} \sqrt{x}$

$$\int_{1}^{\infty} \alpha^{x} / \sqrt{x} \, dx = 2 / \sqrt{-\ln \alpha} \int_{\sqrt{-\ln \alpha}}^{\infty} e^{-t^{2}} \, dt.$$

Also,

$$\int_{\sqrt{-\ln \alpha}}^{\infty} e^{-t^2} dt < \int_{0}^{\infty} e^{-t^2} dt = \sqrt{\pi} / 2$$

and

$$\int_{\sqrt{-\ln \alpha}}^{\infty} e^{-t^2} dt = \int_{0}^{\infty} e^{-t^2} dt - \int_{0}^{\sqrt{-\ln \alpha}} e^{-t^2} dt$$

$$> \int_0^\infty e^{-t^2} dt - \int_0^{\sqrt{-\ln \alpha}} 1 dt$$
$$> \sqrt{\pi} / 2 - \sqrt{-\ln \alpha} .$$

Combining these inequalities we obtain

$$\sqrt{\pi/(-\ln lpha)} - 2 < \sum_{n=1}^{\infty} \alpha^n / \sqrt{n} < \alpha + \sqrt{\pi/(-\ln lpha)}$$

With the values of α and c from section 4, we get

$$1,902,750 < \sum_{n=1}^{\infty} \alpha^n / \sqrt{n} < 1,902,760$$

and

$$104,218 < \sum_{n=1}^{\infty} c \, \alpha^n / \sqrt{n} < 104,219.$$

The bound of Stirling's approximation of factorials [6] reveals that for each n,

$$0.9988c\alpha^n/\sqrt{n} < P(A_n) < c\alpha^n/\sqrt{n}$$

and so

$$104,092 < \sum_{n=1}^{\infty} P(A_n) = a < 104,219.$$

Finally, since $P(\bigcup_{n=1}^{\infty}A_n) = f = a/(1+a)$,

$$0.99999039 < P\left(\bigcup_{n=1}^{\infty} A_n\right) < 0.99999041.$$

We conclude that it is not *certain* that Lazzarini would eventually have obtained the optimum $\hat{\pi}$, but the odds favoring it are overwhelming (assuming no measurement error).

Note: The estimation of the above numerical values and that of α originally caused many headaches. While this work was in progress my school got a multiple precision arithmetic package. The headaches went away! Students' reliance on calculators may be shaken by asking them to calculate α . Some methods yield results less than 1, some greater than 1 and some yield overflow.

6. The Case Against Lazzarini

The result of section 5 suggests that it is at least plausible, ignoring measurement error, that Lazzarini actually performed the experiment. However, Lazzarini did not report just a single experiment of 3408 casts and 1808 hits. He reported a series of casts:

N	100	200	1000	2000	3000	3408	4000
Н	53	107	524	1060	1591	1808	2122

It is highly suspicious that all the values of H in this series are very close to their expected values Np, which are respectively 53.05, 106.10, 530.52, 1061.03, 1591.55, 1808.00, and 2122.07. Ordinarily we expect much greater fluctuations than this in random data. In fact, even if we only look at those hits when N is a multiple of 1000, the probability of being this close to the expected values is exceedingly small. Define G_k to be the event that the number of hits, H_k , in 1000k casts is at least as close to the expected number of hits as Lazzarini reported. Then

$$G_{1} = \{524 \le H_{1} \le 537\}$$

$$G_{2} = \{1060 \le H_{2} \le 1062\}$$

$$G_{3} = \{1591 \le H_{3} \le 1592\}$$

$$G_{4} = \{H_{4} = 2122\} \text{ and}$$

$$P(G_{1} \cap G_{2} \cap G_{3} \cap G_{4})$$

$$= P(G_{1}) \cdot P(G_{2}|G_{1}) \cdot P(G_{3}|G_{1} \cap G_{2}) \cdot P(G_{4}|G_{1} \cap G_{2} \cap G_{3})$$

$$\le P(G_{1}) \cdot P(G_{2}|H_{1} = 531) \cdot P(G_{3}|H_{2} = 1061) \cdot P(H_{4}|H_{3} = 1592)$$

$$\le P(524 \le H_{1} \le 537) \cdot P(1060 - 531 \le H_{1} \le 1062 - 531)$$

$$\cdot P(1591 - 1061 \le H_{1} \le 1592 - 1061) \cdot P(H_{1} = 2122 - 1592).$$

Using the normal approximation to the binomial probabilities we obtain a probability of less than 0.00003. Thus it seems exceedingly unlikely that Lazzarini carried out a random series of tosses with results as nearly optimal as he reported. So it seems likely the experiment was not done—at least not in a random fashion.

7. Speaking of Hoaxes

But setting aside measurement error and granting that the experiment was a hoax, one may ponder the quality of the hoax. Here are three hoaxes to compare with that of Lazzarini.

In hoax one, let d = 10 cm and l = 7.1 cm. It seems not unlikely that a garden variety needle might measure 7.1 cm and a round figure for d > l is d = 10 cm. In this hoax any multiple of 250 casts is a potential generator of 355/113—one needs the same multiple of 113 hits. This hoax has the advantage that the optimal stopping points are plausible; for instance, every multiple of 1000 is such a point.

In hoax two, let l = 7.1 cm and d = 11.3 cm. These seem like less plausible "objective" values of l and d, but in the experiment, the occurrence of 355/113 is more assured because every multiple of five casts is a potential generator.

In hoar three, we go for an even more accurate $\hat{\pi}$. As mentioned earlier, the next improvement on 355/113 is 52,163/16,604. This was found by computer and may not have been known to nineteenth-century mathematicians. But continued fraction convergents were available and the next convergent after 355/113 is 103,993/33,102. But 103,993 is a prime and so this convergent is not a good candidate, but the one after that, 104,348/33,215, does lead to a plausible hoar. Since $104,348 = 2 \cdot 2 \cdot 19 \cdot 1373$ and $33,215 = 5 \cdot 7 \cdot 13 \cdot 73$, we can take $l = 2 \cdot 2 \cdot 1373 \times 10^{-3} = 5.492$ cm, and $d = 7 \cdot 13 \cdot 73 \times 10^{-3} = 6.643$ cm. These aren't particularly plausible but they are within the accuracy that was then measurable and the payoff is that every multiple of 19 casts is a potential generator of $\hat{\pi}$ that misses π by less than 3×10^{-10} and so is accurate in the ninth decimal place.

The advantage of hindsight (which Lazzarini lacked) allows us to design bogus experiments that foil today's statistical tests designed to expose them. It seems to me that hoax one does just that. This is especially true if only a final value of $\hat{\pi}$ is reported or, if reported, a series of values of hits has sufficiently random dispersion.

Today, one occasionally hears of bogus experiments and/or rigged data [2], [10]. Presumably modern hoaxers are aware of the type of statistics that exposed the Lazzarini hoax and put enough dispersion in their "data" to avoid the same fate. It will be interesting to see what, if any, (future) statistical or scientific test will be brought to bear on their work, thereby labeling them as only poor hoaxers and properly removing them from the ranks of objective scientists.

Acknowledgements. I want to thank the referees for their helpful suggestions and support. I especially want to thank my friend and colleague, Jim Foster. Jim's input helped convert my qualitative findings into quantitative probabilities and accounts for a good deal of the precise character of the results.

REFERENCES

- 1. P. Beckmann, A History of π , The Golem Press, New York, 1971.
- 2. W. Broad and N. Wade, Betrayers of the Truth, Simon and Schuster, New York, 1982.
- 3. J. Bronowski, The Common Sense of Science, Harvard University Press, Cambridge, 1953.
- 4. G. L., C. Buffon, Essai d'arithmétique morale, in Vol. 4 of the Supplément à l'Histoire Naturelle, 1777.
- 5. H. Eves, An Introduction to the History of Mathematics, CBS College Publishing, New York, 1983.
- W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 1, John Wiley & Sons, Inc., New York, 1957.
- 7. G. Gamow, One Two Three · · · Infinity, The Viking Press, New York, 1961.
- 8. B. V. Gnedenko, The Theory of Probability, 4th ed, (English translation), Chelsea Publishing Co., New York, 1967.
- 9. N. T. Gridgeman, Geometric probability and the number π , Scripta Math., 25 (1960), 183–95.
- 10. D. P. Hamilton, Verdict in sight in the "Baltimore Case", Science, Vol. 251 (8 March 1991), 1168-72.
- 11. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, London, 1979.
- 12. IBM, Precision Measurement in the Metal Working Industry, Syracuse U. Press, Syracuse, NY, 1952.
- 13. E. Kasner and J. Newman, Mathematics and the Imagination, Simon and Schuster, New York, 1940.
- 14. P-S. Laplace, Théorie Analytique des Probabilités, 1812.
- 15. M. Lazzarini, Un' applicazione del calcolo della probabilità alla ricerca sperimentale di un valore approssimato di π , *Periodico di Matematica*, 4 (1901), 140–3.
- 16. T. H. O'Beirne, Puzzles and Paradoxes, Oxford University Press, London, 1965.
- 17. B. Pilton, Genuine fakes, in Seven Years of Manifold, ed. by Ian Stewart and John Jaworski, Shiva Publications, Nantwich, 1981.
- 18. A. N. Zaydel, Delusion or fraud, Quantum, September/October 1990, 6-9.

WEATHER REPORT ON T.V. NEWS, 11:15 p.m. November 21, 1990 (reported by Ralph Boas)

"It's warmer right now than today's high."